

# Does Sex Speed Up Evolutionary Rate and Increase Biodiversity?

Carlos J. Melián<sup>1,2\*</sup>, David Alonso<sup>3</sup>, Stefano Allesina<sup>1,4</sup>, Richard S. Condit<sup>5</sup>, Rampal S. Etienne<sup>3</sup>

**1** National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, California, United States of America, **2** Center for Ecology, Evolution and Biogeochemistry, Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland, **3** Community and Conservation Ecology Group, University of Groningen, Groningen, The Netherlands, **4** Department of Ecology and Evolution, University of Chicago, Chicago, Illinois, United States of America, **5** Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Balboa, Ancon, Republic of Panama

## Abstract

Most empirical and theoretical studies have shown that sex increases the rate of evolution, although evidence of sex constraining genomic and epigenetic variation and slowing down evolution also exists. Faster rates with sex have been attributed to new gene combinations, removal of deleterious mutations, and adaptation to heterogeneous environments. Slower rates with sex have been attributed to removal of major genetic rearrangements, the cost of finding a mate, vulnerability to predation, and exposure to sexually transmitted diseases. Whether sex speeds or slows evolution, the connection between reproductive mode, the evolutionary rate, and species diversity remains largely unexplored. Here we present a spatially explicit model of ecological and evolutionary dynamics based on DNA sequence change to study the connection between mutation, speciation, and the resulting biodiversity in sexual and asexual populations. We show that faster speciation can decrease the abundance of newly formed species and thus decrease long-term biodiversity. In this way, sex can reduce diversity relative to asexual populations, because it leads to a higher rate of production of new species, but with lower abundances. Our results show that reproductive mode and the mechanisms underlying it can alter the link between mutation, evolutionary rate, speciation and biodiversity and we suggest that a high rate of evolution may not be required to yield high biodiversity.

**Citation:** Melián CJ, Alonso D, Allesina S, Condit RS, Etienne RS (2012) Does Sex Speed Up Evolutionary Rate and Increase Biodiversity? PLoS Comput Biol 8(3): e1002414. doi:10.1371/journal.pcbi.1002414

**Editor:** Jan Engelstädter, ETH Zurich, Switzerland

**Received:** August 4, 2011; **Accepted:** January 20, 2012; **Published:** March 8, 2012

**Copyright:** © 2012 Melián et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** CJM, SA, and RSC were supported by fellowships from the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant \#DEB-0553768), the University of California, Santa Barbara, and the State of California. CJM also acknowledges the support by Microsoft Research Ltd., Cambridge, United Kingdom. DA and RSE acknowledge the support of the Netherlands Organization for Scientific Research (NWO). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: carlos.melian@eawag.ch

## Introduction

The impact of sexual reproduction on the rate of evolution could stand as one of biology's grand achievements [1–4]. Does sex speed genetic divergence, speciation, and thus increase the world's diversity relative to asexual reproduction? An immediate difficulty with any theory is how to define speciation in asexual organisms, where Mayr's Biological Species Concept [5] does not easily apply [6,7]. Nevertheless, asexual organisms do diversify and are assigned species names [8–12], and many observations and experiments describe speciation in sexual as well as asexual organisms. Much work emphasizes ecological divergence and speciation [13–17], but we propose to step back and ask basic questions about the dynamics of divergence and extinction, and how it depends on sexual reproduction. Before we understand the full impact of sex on evolution and diversity in an ecologically complex world, we need to understand well the basic dynamics of mutation, gene flow, drift and extinction underlying the process of speciation.

Sex increases the rate of evolution [18–24], although evidence of sex constraining genomic and epigenetic variation and slowing down evolution also exists [25–27]. Given these contrasting impacts of sex, the effects of reproduction mode on patterns of

diversification, extinction and consequent species diversity are hard to predict, even without ecological opportunity. We here pose a basic question to connect the dynamics of sexual and asexual populations with biodiversity patterns: do sexually reproducing populations have similar biodiversity patterns as asexual populations in the absence of ecological differentiation, given equal mutation and identical definitions of the genetic divergence required for speciation? How do mutation, genetic drift, ecological drift, and gene flow act in sexual, versus asexual, populations to produce diversity?

Research on diversification of species often emphasizes the process of genetic divergence, but extinction rates are also critical. Even in the absence of selection, the dynamics of diversification and diversity may thus be influenced by mutation, genetic drift, sexual recombination, colonization, as well as population size and its role in ecological drift and extinction [17]. Other than the direct impact of sexual recombination on genetic divergence, are other aspects of the dynamics of evolution the same in sexual and asexual populations? We take here a theoretical approach to the genetics of speciation [28–32] in the context of neutral biodiversity theory [33]. Our goal is to model the emergence of new species using explicit genetic rules on a backdrop of individuals whose births and deaths determine abundances and extinction. This

## Author Summary

The role of sex in driving genetic variation and the speed at which new species emerge has been debated for over a century. There is experimental and theoretical evidence that sex increases genetic variation and the speed at which new species emerge, although evidence that sex reduces variation and slows the formation of new species also exists. Surprisingly, given the link between sex and genetic variation, little work has been done on the impact of sex on biodiversity. In the present theoretical study we show that a faster evolutionary rate can decrease the abundance of newly formed species and thus decrease long-term biodiversity. This leads to the paradoxical result that sexual reproduction can increase genetic variation but reduce species diversity. These results suggest that reducing the rate of appearance of genetic variation and the speed at which new species emerge may increase biodiversity in the long-term. This unexpected link between reproductive mode, the speed of evolution and biodiversity suggests that a high evolutionary rate may not be required to yield a large number of species in natural ecosystems.

genetic model of speciation extends existing neutral models of community diversity [34–44] so that speciation, extinction, abundance, the population size of new species, and diversity emerge from assumptions on genetic divergence, and genetic and ecological drift. Modeling speciation as a neutral process is unrealistic, but this simplification may serve as a useful null model to compare speciation rate and species diversity in asexual vs. sexual communities, minimizing the number of assumptions about population-level patterns of speciation and extinction.

To understand the effect of reproductive mode on patterns of diversification and species diversity, we need a model describing the dynamics of genes within populations within a model of populations within a community, and we need a definition of speciation that applies to sexual and asexual populations [8,45]. Our definition of species is used in the context of a population whose genomes diverge in a spatial landscape. In the model, a community without deme structure and no ecological differentiation [46] has  $N_e$  individuals, and the geographic distance between each pair of individuals  $i$  and  $j$  is given by  $d_{ij}$ ;  $D$  is the geographic distance matrix containing all the  $d_{ij}$  values. All individuals have identical and (essentially) infinite genomes of  $L$  nucleotides at the outset (see “Material and Methods” and “Table 1” for a summary of the mathematical terms used in this analysis). The genetic similarity,  $q$ , between each pair of individuals  $i$  and  $j$  can be represented by a genetic similarity matrix,  $Q = [q_{ij}]$ . At time 0, all elements are  $q = 1.0$ , but there is a constant mutation rate  $\mu$  per nucleotide per birth-death cycle, so the community evolves divergence under the combined influences of recombination in sexual populations or asexual reproduction, but not both, mutation, migration, and genetic drift. We assume asexual individuals are strictly asexuals, with no horizontal gene transfer. In sexual populations, pairs mate and exchange sections of their genomes. In both models, dispersal and colonization is incorporated because offspring appear near their parents, and mating is among neighbors.

The crucial, final feature needed to make this a model of speciation is a minimum genetic similarity threshold,  $q^{\min}$  [47,48]: two individuals  $i$  and  $j$  for which  $q_{ij} < q^{\min}$  are sufficiently different to be called different species. In the sexual population, this means those two individuals cannot mate; in the asexual community, it has no impact on dynamics for the obvious reason that there is no

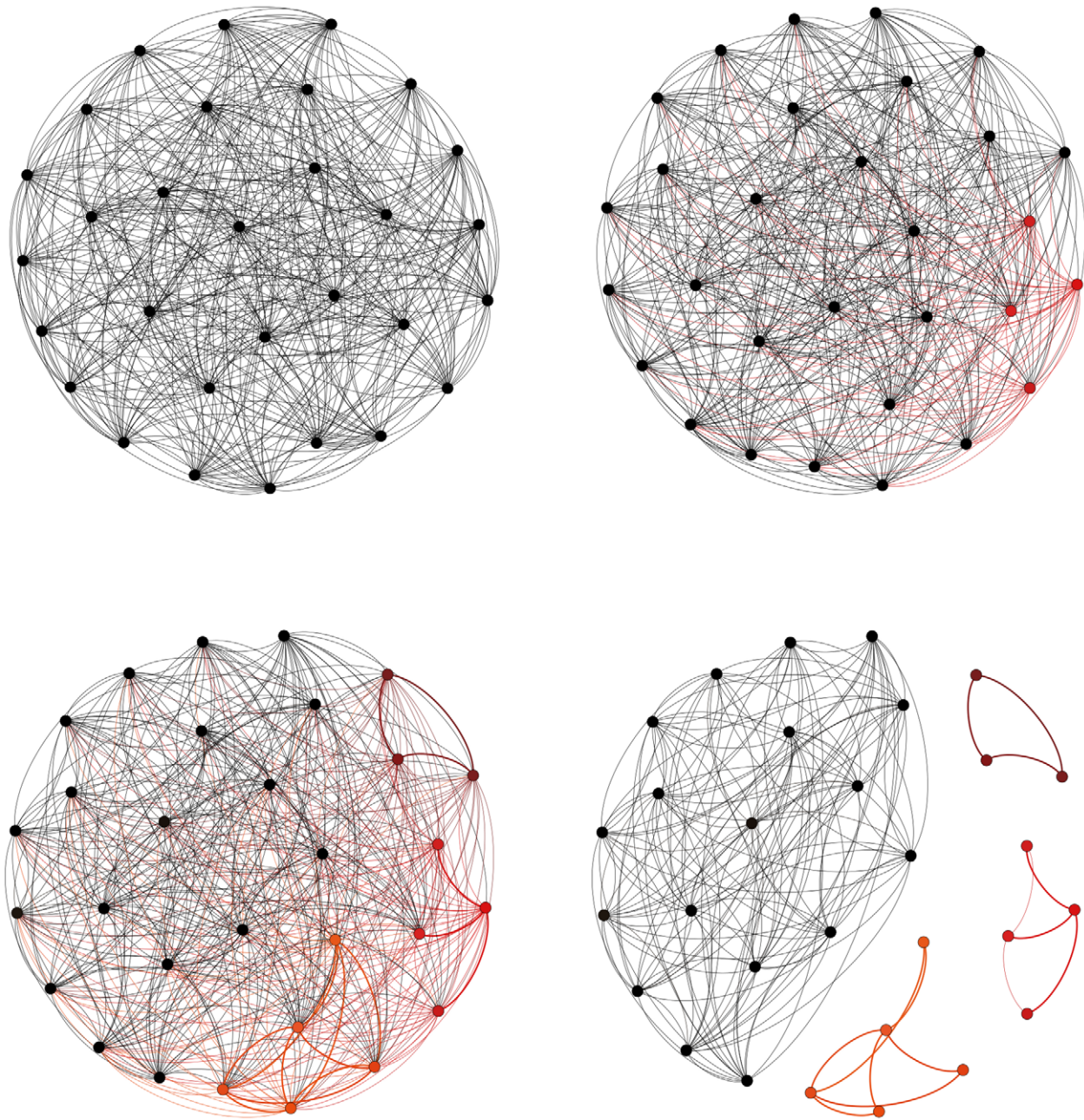
**Table 1.** Glossary of mathematical notation.

Notation	Definition
$d_{ij}$	Geographical distance between individual $i$ and $j$
$D$	Geographic distance matrix containing all the $d_{ij}$ values
$d^{\max}$	Maximum geographical distance to find a mating partner and dispersal
$q_{ij}$	Genetic similarity between individual $i$ and $j$
$Q$	Genetic similarity matrix containing all the pairwise similarity $q_{ij}$ values
$q^{\min}$	Minimum genetic similarity above which $i$ and $j$ belong to the same species
$\bar{Q}$	Mean genetic similarity of the matrix $Q$ in the transient
$Q^*$	Expected mean genetic similarity of matrix $Q$ at equilibrium
$\mu$	Mutation rate per nucleotide per birth-death cycle
$N_e$	Effective population size
$L$	Length of the genome
$E[q^{kj}]$	The expected genetic similarity between the new offspring $k$ and each individual $j$ in the population
$f^{ij}$	The fraction of identical sites between individual $i$ and $j$
$S_u^i$	The $u^{\text{th}}$ site in the genome of individual $i$
$n_{\text{sex}}^*$	The number of direct links in a chain of inheritance before ancestor and descendant are more different than the genetic cut-off of species formation in sexual populations
$n_{\text{asex}}^*$	The number of direct links in a chain of inheritance before ancestor and descendant are more different than the genetic cut-off of species formation in asexual populations

doi:10.1371/journal.pcbi.1002414.t001

mating. In both cases, we imagine that a biologist observing these two individuals would be inclined to describe them as different species; in the sexual case, the same biologist would detect sufficient genetic incompatibilities that offspring would be inviable [49–53,53,54]. Defining critical divergence for a pair of individuals, however, is not yet a species definition, because species boundaries are a property of entire populations [31,53,55]. A species is defined as a connected component in an evolutionary graph: a group of individuals for which there is a path of genetic compatibility connecting every pair (Fig. 1). This means that two individuals in sexual populations can be conspecific while also being incompatible, as long as they can exchange genes indirectly through other conspecifics (a ring species [56]). Using this definition, speciation will occur whenever the expected mean genetic similarity of the matrix  $Q$  at equilibrium reaches  $Q^* < q^{\min}$  [28,57]; intuitively, this is straightforward: with mutation too low, so  $Q^* > q^{\min}$ , the community reaches an equilibrium similarity,  $Q^*$ , so speciation does not start (pp. 305, [57]).

In summary, the process of diversification starts with an initial phase during which genetic similarity gradually declines toward an equilibrium,  $Q^*$ . Individuals become more and more divergent from one another, particular those further away in space.



**Figure 1. Diversification in spatial networks.** **Top left,** In the initial stage all individuals, represented as black nodes, are reproductively compatible corresponding to a completely connected graph. At this stage, distance edges, represented by the geographic distance matrix,  $D$ , containing all the  $d_{ij}$  values, capture both geographical separation of each pair of individuals and viable edges. **Top right and bottom left,** Species formation start in the transients (red circles, top right and dark red, red and orange, bottom left). A species is defined as a population whose genetic similarity of each pair of individuals within the population is above a minimum genetic similarity threshold,  $q^{\min}$ . For example, the genetic similarity between each pair of individuals  $i$  and  $j$ ,  $q_{ij}$ , within the population in red satisfies  $q_{ij} > q^{\min}$ . At this stage all the individuals in the network are still reproductively compatible. Formed species have different abundance (i.e, dark red (3), red (4) and orange (5)). **Bottom right,** In the last stage, individuals within each species are reproductively isolated to all other individuals in the population. For example, each pair of individuals  $i$  and  $j$  within the species in red now satisfy  $q_{ij} > q^{\min}$  and  $q_{ik} \leq q^{\min}$  for all the individuals  $k$  outside the population.  
doi:10.1371/journal.pcbi.1002414.g001

Eventually, two clusters are formed with the special property that there is not a single individual in one that is compatible with any individual in the second: they form two species (Fig. 1). This is permanent, for once segregated, because given a very large genome with  $L$  nucleotides, the universe of possible genome configurations is so large relatively to population size that it is essentially impossible for compatibility of genomes to become

reestablished. The divergence process continues until each of those clusters divides further, and so on. Each speciation event leads to a loss of divergence within species, followed again by increasing divergence until another speciation event. Thus, genetic similarity within species is blocked from ever falling (much) below  $q^{\min}$  by the speciation process. But there is still more to the dynamics of diversification, due to extinction. Once we assume species



formation, we must include ecological drift – random fluctuations of species abundances within a fixed metacommunity size – as an influence on dynamics, in addition to reproductive mode, mutation, migration and genetic drift. Species may start rare, or become rare due to drift, and then go extinct, and speciation should eventually be balanced by extinction, exactly as the neutral model of diversity describes [33].

We now examine this model in detail to ask whether sexual or asexual populations (and metacommunities) give rise to faster diversification or more species, considering the equilibrium, at which speciation and extinction have reached a balance, as well as the transient increase in diversity after a founding event. First, a theoretical analysis of the divergence process leads to important assertions about the speciation rate and how it relates to the mutation rate in sexual versus asexual communities. The full spatial model in which ecological drift controls the equilibrium diversity requires simulations, and we ran models with a wide variety of parameter combinations in order to answer the main questions: 1) Do species appear faster in a sexual or an asexual population? 2) Is species richness higher at equilibrium in sexual or asexual communities? 3) Are species abundances different, so does ecological drift play an important role in the extinction rate?

## Results

We first consider analytically the time course of differentiation and speciation by considering the number of steps in a chain of descendants until the threshold of genetic divergence is reached. That is the stage at which a descendent is incompatible with a founding ancestor: speciation can only happen after that. Consider asexual populations and examine one individual  $A$  and its descendants. Denote successive individuals  $A \rightarrow A_1 \rightarrow A_2, \dots, A_n$ , where  $A_{i+1}$  is the offspring of  $A_i$ . We determine the number of steps until a descendant is sufficiently different from  $A$  to be incompatible, so  $q^{AA_n} < q^{\min}$ . Genetic similarity between  $A$  and  $A_1$  after the first offspring (following equation 18 in “Material and Methods”) is

$$q^{AA_1} = x_1 = e^{-2\mu}, \quad (1)$$

where  $\mu$  is mutation rate; after  $n$  offspring, it is

$$q^{AA_n} = x_n = e^{(-2\mu)n}. \quad (2)$$

The critical step, where  $q^{AA_n} < q^{\min}$ , is therefore

$$n_{\text{asex}}^* = -\frac{\log(q^{\min})}{2\mu}, \quad (3)$$

A curious mutation rate is the one which produces a new cluster or species after the first offspring, so  $n^* = 1$ , a rate so high that offspring are “hopeful monsters”: different species from their parents [58]. We call this “mutation-induced” speciation, but such a model makes little biological sense.

The model produces species with a different mechanism with much lower mutation rates: “fission-induced” speciation. Imagine a chain of descendants  $A \rightarrow A_1 \rightarrow A_2, \dots, A_n$  in which every individual is alive. Ignoring the “hopeful monsters” of mutation-induced speciation, the entire chain belongs to the same cluster or species, even if  $A$  and  $A_n$  are distinct enough to be incompatible. After enough time, however, intermediate steps in the chain die, and eventually a subchain  $A_i \rightarrow A_{i+1} \rightarrow A_{i+2}, \dots, A_{i+m}$  is entirely

dead. Once a subchain of  $m$  consecutive steps, with  $m > n^*$ , dies, the survivors in the chain become two separate species. Obviously, at some point there is a single critical individual whose death breaks the single cluster into two clusters – the last of the  $m$  individuals in the subchain to die.

With fissioning of genetic clusters, new species need not be singletons. Indeed, there is no upper limit on the abundance of a new species (the parent population size is the upper limit). Incipient population size should depend on  $n^*$ , and thus the minimum genetic similarity value,  $q^{\min}$ , that defines a species, and mutation rate ( $\mu$ ): the higher  $n^*$ , the more time it will take before we have a new cluster formation. With higher  $n^*$ , we thus anticipate lower speciation rate, but lineages may have higher incipient abundances and thus be less prone to extinction.

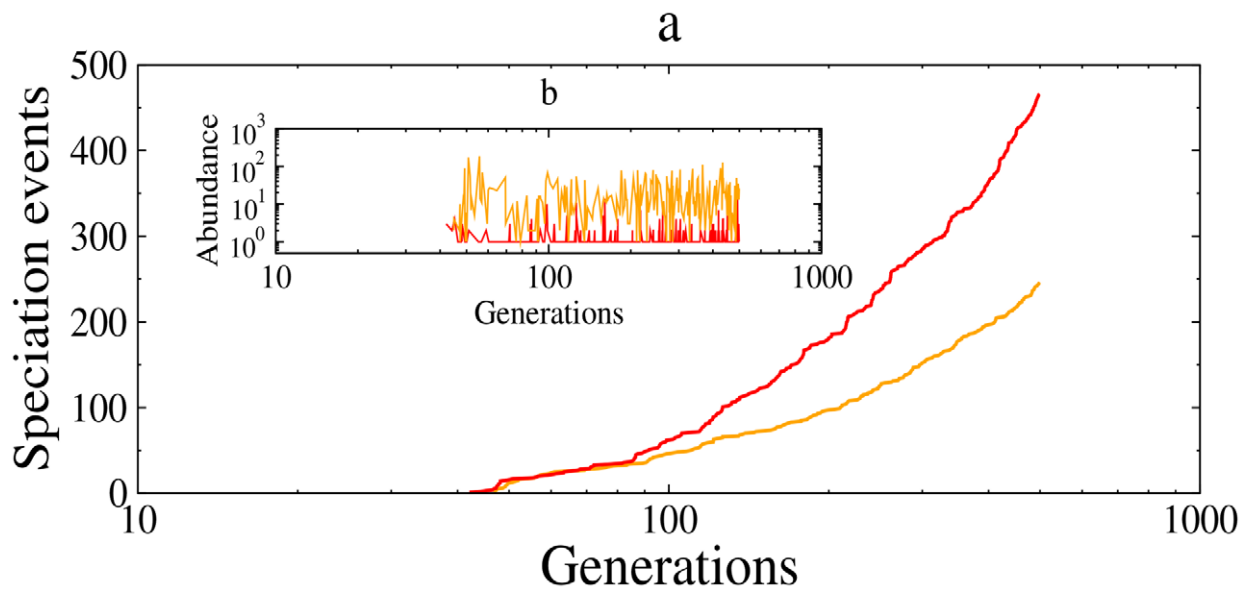
In an earlier work, we examined the dynamics of the number,  $n^*$ , in sexual populations in the absence of a limited geographical distance for mating ( $d^{\max} = 1$ ) [32]. The critical number of steps where  $q^{AA_n} < q^{\min}$  in a panmictic population with sexual reproduction is:

$$n_{\text{sex}}^* = \frac{\log(q^{\min})}{-2m + \log[(q^{\min} + 3)/4]}, \quad (4)$$

The extra  $\log$  term in the denominator compared to equation 3 reflects genetic difference between mated pairs, and thus genetic dissimilarity between offspring and parents beyond mutation. Equations 3 and 4 suggest that there should be a monotonic relationship between  $n^*$  and the speciation rate. New species will form at the rate at which chains of length  $n^*$  die. Comparing equations 3 and 4, we observe that  $n_{\text{asex}}^* > n_{\text{sex}}^*$  in all cases, so  $1/n_{\text{sex}}^* > 1/n_{\text{asex}}^*$ . With sex, it takes fewer steps,  $n^*$ , before a descendent passes the critical genetic similarity,  $q^{\min}$ , relative to its ancestor and this should lead to a lower speciation rate in an asexual metacommunity at a given mutation rate.

Simulations confirm this assertion. Soon after founding, the metacommunity with sexual reproduction produces species more rapidly at a given mutation rate,  $\mu$ , than the asexual case (Fig. 2a), but lineages have lower abundances and thus are more prone to extinction (Fig. 2b). This pattern has strong consequences for species richness (Fig. 3). In the transient, at very high mutation rate, the number of species collapsed in both models (Fig. 3b). At equilibrium, quite surprisingly, the opposite held, and the asexual model had higher number of species for low mutation rate values (Fig. 3c). The sexual model was much less efficient at maintaining species despite the higher rate of species formation. At  $\mu \in \{10^{-4}, 10^{-5}\}$ , for instance, there are 1–3 species in simulations with sex, compared to 2–10 species without sex (Fig. 3c). These patterns remain the same after we compare the transient (Fig. 3d) and the steady-state (Fig. 3e) regardless of the maximum geographic distance for mating and dispersal,  $d^{\max}$ .

This failure to maintain species richness in the sexual model could only have been due to extinction: at a given diversity, the sexual communities must have lost species at a higher rate than asexual communities. This would happen if incipient species abundances were more skewed in the sexual model. Fig. 4 shows that this is indeed the case. In sexual communities, there were more incipient species with low abundance in the transient (red line, Fig. 4a) and at equilibrium (red line, Fig. 4b), very few highly abundant species, and many rare species (Fig. 4c) relative to asexual communities. This pattern remains qualitatively the same in all pairwise comparisons between sexual and asexual metacommunities with  $\mu < 10^{-4}$  and  $d^{\max} \in \{1, 0\}$  (Kolmogorov-Smirnov test,  $p < 0.001$ ).



**Figure 2. Speciation and incipient species abundance.** **a**, Cumulative number of speciation events as a function of the number of generations for the sexual (red, also used for b) and asexual populations (orange also used for b). **b**, Simulated abundance of the new species after each speciation event. The plot represents the output from one replicate during the first 500 generations with mutation rate,  $\mu = 10^{-4}$ , the minimum genetic similarity value,  $q^{\min} = 0.97 > Q^*$ , and a maximum geographic distance for mating and dispersal,  $d^{\max} = 1$ . doi:10.1371/journal.pcbi.1002414.g002

## Discussion

In the present paper, we have explored a landscape population genetics model to understand the effect of reproductive mode on speciation and extinction rate and the connection between the abundance of new species and species richness. The approach uses processes of individual organisms with large genomes – birth, death, gene flow, mutation and genetic-ecological drift – to study macroecological patterns of biodiversity [28–32]. It allows a comparison of diversification rate and community diversity in sexual vs. asexual communities without recourse to any assumption about population-level patterns of speciation and extinction. By modeling speciation explicitly, genetic assumptions about the formation of species become necessary: in the present study, the constant mutation rate and threshold of genetic similarity defining the species boundary.

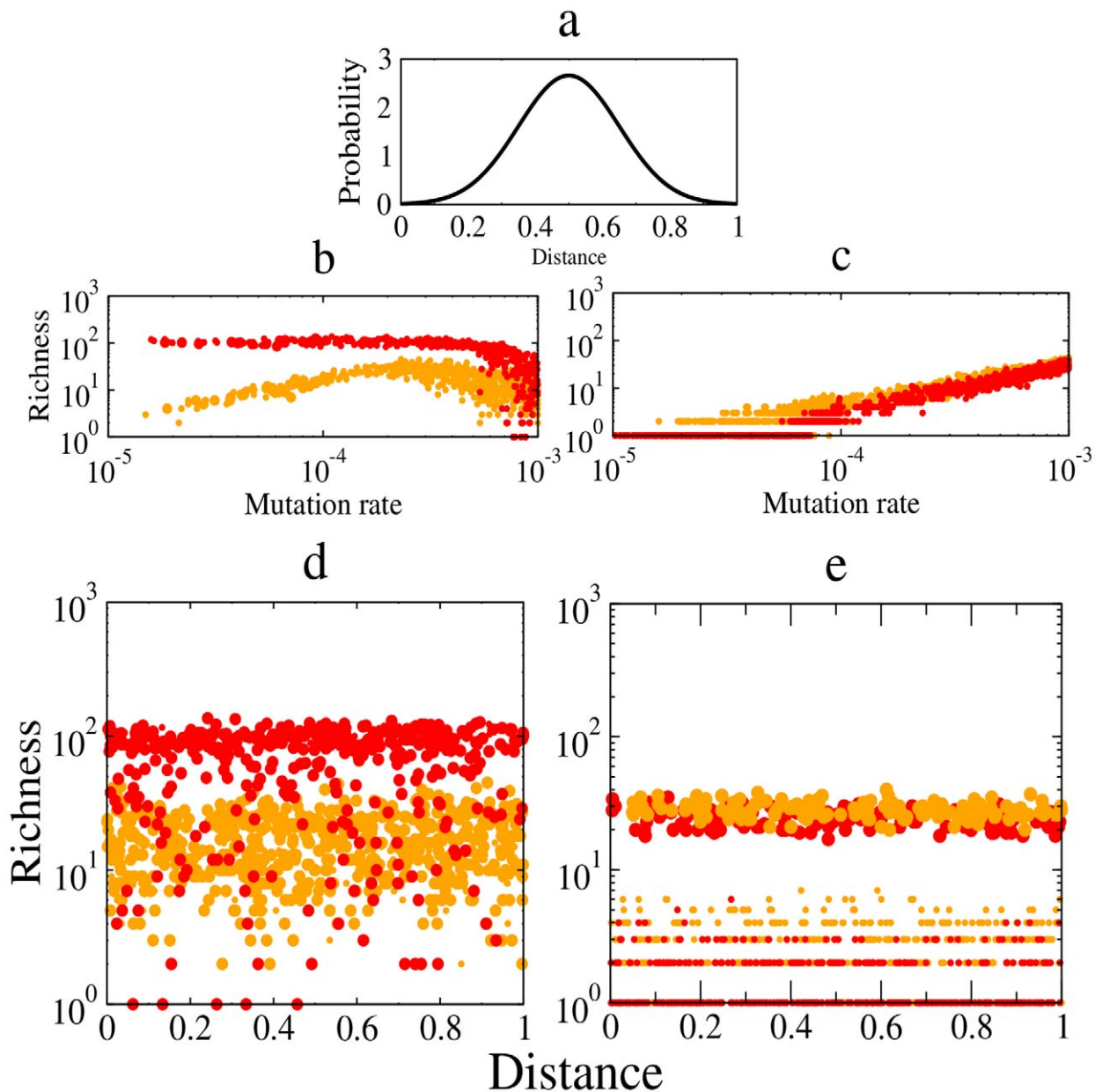
These assumptions allow us to derive quantitative relationships between mutation rate, abundances, probability of extinction of new species, and species richness. For example, the number of species in a metacommunity increases monotonically with mutation rate in both sexual and asexual populations. But mutation alone cannot cause speciation, because the genetic similarity defining species is also essential, entering in the equations that drive the rate of species formation (i.e.,  $1/n_{\text{asex}}^*$  and  $1/n_{\text{sex}}^*$ ). The quantitative nature of the relationship between mutation rate, genetic similarity, and species formation is understood with  $n^*$ : incipient population size and species richness are both functions of  $n^*$ , and higher  $n^*$  means more time between speciation events but also higher incipient abundances and lower extinction rate.

Surprisingly, sexual populations, with low  $n^*$  value and thus a high speciation rate, had greatly reduced species richness at equilibrium, relative to the asexual populations with otherwise similar processes. This highlights the importance of deriving the processes connecting the rate of evolution and incipient abundance – the number of individuals in newly formed clusters or species – because they both impact speciation [35,59] and the

number of species that can coexist in metacommunities. Incipient species abundance was highly variable in both sexual and asexual populations, but especially so with sex. In the latter, newly formed species were often singletons and thus rapidly went to extinction.

Most speciation events in nature are believed to have been driven by divergent selection and drift is thought to play a very small role [15,53]. But genetic and ecological drift can be strong contributors during speciation, especially in the early stages [17]. We have shown that a higher evolutionary rate in sexual populations does not guarantee more coexisting species, especially in the long term, because higher evolutionary rate may imply a lower abundance of new species and thus a higher extinction probability. Thus, even if drift plays a small role in driving differentiation and speciation, it can strongly influence the extinction dynamics driven by the low abundance of the incipient species in natural populations.

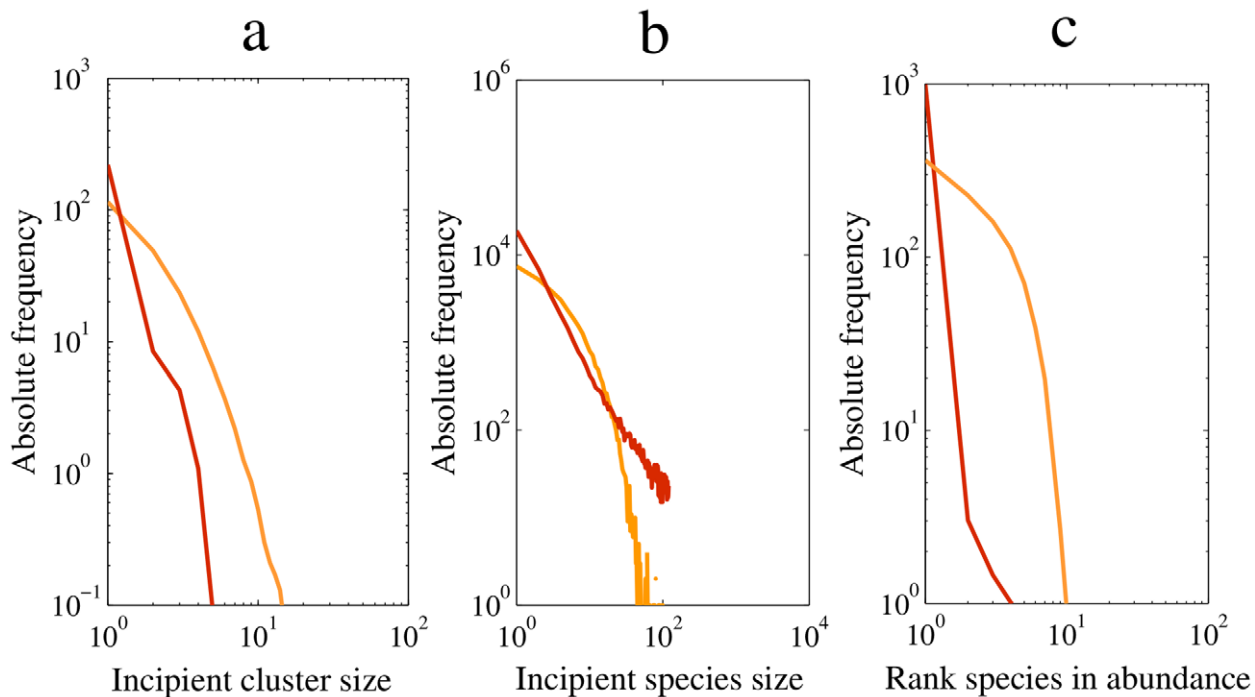
How robust are these results after the addition of selection and ecological differentiation? Sexual organisms might have faster rates of adapting to different ecological conditions [21–24,60], because multiple beneficial mutations can spread simultaneously in the population [20]. This can trigger higher abundance and lower extinction probability in sexual populations because speciation is not being driven by mutation but rather by adaptations to ecological conditions. Sex can also constrain the rate of adaptation to new conditions [25,26]. For example, it removes major changes such as chromosomal rearrangements [27], and in the process of finding a mate, it may increase the risk to predation or higher exposition to sexually transmitted diseases [61,62]. These processes may slow down the rates of evolution and speciation which, according to our results, may not necessarily decrease the number of species in the long term. Further research that connects genetic and ecological drift with selection in constant and fluctuating environments may shed light on the link between reproductive mode, the rate of speciation, the abundance of new species, extinction probability, and long term species richness [17,63].



**Figure 3. Speciation, extinction and biodiversity in spatial networks.** **a**, Geographic distance between individual  $i$  and  $j$ ,  $d_{ij}$ , is sampled from a normal distribution,  $\mathcal{N}(\bar{d}, s^2)$ . In this plot the mean,  $\bar{d}=0.5$ , and the standard deviation,  $\sigma^2=0.35$ . **b–c**, Species richness as a function of mutation rate,  $\mu$ , with maximum geographic distance values to find a mating partner and dispersal in the range,  $d^{\max} \in \{0, 1\}$  for asexual (orange) and sexual (red) populations in the transient (**b**) and last (**c**) stage. **d–e**, Species richness in the transient (**d**) and last (**e**) stage as a function of the maximum geographic distance to find a mating partner and dispersal,  $d^{\max}$  ("Distance"). The size of the circles represent the species richness at a given mutation rate in the range,  $\mu$  (big circles)  $\in \{8 \times 10^{-5}, 10^{-3}\}$  and  $\mu$  (small circles)  $\in \{8.1 \times 10^{-5}, 10^{-5}\}$ . In this plot each replicate in the transients (**d**) satisfies,  $q^{\min} = 0.98 > \bar{Q} = 0.97$ . Sexual populations have more species in the transients for a broad range of parameter values. Richness collapses for high mutation rate values (see **b**), thus the high dispersion of species richness values in **d**. Asexual populations maintain higher biodiversity levels than the sexual populations, especially for low mutation rate values (see **c** and **e**). Each replicate in the last stage (**c** and **e**) satisfies  $q^{\min} \in \{0.97, 0.90\} > \bar{Q}^*$ .  
doi:10.1371/journal.pcbi.1002414.g003

A theory that covers the link between net diversification rates and biodiversity patterns in both sexual and asexual taxa is still lacking [17,21,64,65], and our approach joining population genetics models with divergence criteria and macroecological patterns of biodiversity may be a way forward. One important advance would be to develop analytical relationships among the key parameters, particular mutation rate and the strength of selection in the context of several topologies of spatial networks, and subsequently spatial

heterogeneity [29,42,57,66,67]. On the genetic side, more precise consideration of the mechanisms driving genome evolution, specifically in the context of rates of self-compatibility and outcrossing, might lead to different predictions about speciation and diversity [10,68–72]. We believe that our results may help to connect reproductive modes with the speciation rate in eco-evolutionary graphs, and the effect of the incipient species abundance on net diversification rate and extant diversity.



**Figure 4. Incipient species abundance and biodiversity.** **a**, Mean incipient species size distribution after  $10^4$  replicates in the first stage for sexual (red) and asexual (orange) populations at a given mutation,  $\mu = 10^{-4}$ . Each replicate satisfies,  $q^{\min} = 0.98 > \bar{Q} = 0.97$ , and a maximum geographic distance for mating and dispersal,  $d^{\max} = 0.5$ . **b**, Mean incipient species size distribution after  $10^4$  replicates in the last stage for sexual (red) and asexual (orange) populations at a given mutation,  $\mu = 10^{-4}$ . Each replicate satisfies  $q^{\min} = 0.92 > \bar{Q}^*$  and a maximum geographic distance for mating,  $d^{\max} = 0.5$ , and **c**, Mean species abundance distribution for sexual (red), and asexual (orange) populations at a given mutation rate,  $\mu = 10^{-4}$ , minimum genetic similarity,  $q^{\min} = 0.90 > \bar{Q}^*$ , and a maximum geographic distance for mating and dispersal,  $d^{\max} = 1$ . doi:10.1371/journal.pcbi.1002414.g004

## Materials and Methods

### The models

Models of DNA evolution based on simple base pair substitution have a long history (i.e., the infinite sites model, [73,74]), and several variants have been proposed [75]. More realistic extensions of those models include deletion, insertion, duplication and rearrangements of segments bases [70]. Recent models also take into account, as in the neutral theory of biodiversity [33], instantaneous speciation but with explicitly evolving genomes (i.e., an identical copy of one root genome is made, each of the two genomes gets a new successor species name, and they each evolve independently thereafter, see [70]).

In the models explored, the reproductive mode describes a population with evolving genomes. During asexual reproduction a mother is randomly selected while in the sexual populations, in addition to this randomly selected mother, potential mates are identified from among those within the specified geographic distance,  $d^{\max}$ . In case there are no potential mates the mother reproduces without a mate. This situation is especially relevant for the extreme case,  $d^{\max} = 0$ . In the sexual and asexual models the offspring is then dispersed within the geographic distance,  $d^{\max}$ , and occupies the site of a randomly killed individual within the area  $d^{\max}$ . At the beginning of the simulation, all individuals are reproductively compatible, corresponding to a completely connected graph (Fig. 1). Genetic similarity among individuals in the sexual and asexual model can be represented by an evolutionary spatial graph in which nodes are individuals, distance edges capture the geographical separation of each pair of individuals and viable edges that connect individuals within the same species.

We here describe formally the derivation of equation 1 in the main text for the asexual model. The dynamics of sexual populations in the absence of dispersal limitation ( $d^{\max} = 1$ ) has been considered elsewhere and will not be derived here [32]. Individuals are haploid. The genome of each individual is represented by a sequence of  $L$  sites, each nucleotide residing in one of two states,  $-1$  or  $+1$ . Each individual  $i$  in a population of size  $J$  is represented as a vector:  $(S_1^i, S_2^i, \dots, S_L^i)$ , where  $S_u^i$  is the  $u^{\text{th}}$  site in the genome of individual  $i$ . The genetic similarity between individual  $i$  and individual  $j$  can be defined as:

$$q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j, \quad (5)$$

with  $q^{ij} \in [-1, 1]$ . The genetic similarity in equation (5) can be written in terms of the fraction of identical sites ( $f^{ij}$ )

$$q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2f^{ij} - 1. \quad (6)$$

and  $f^{ij}$  is:

$$f^{ij} = \frac{1 + q^{ij}}{2}. \quad (7)$$

Each nucleotide in the offspring is inherited at random, thus ignoring linkage between neighboring nucleotides, but with a small

probability of error determined by the mutation rate. Assuming that the individual  $k$  inherited the nucleotide at site  $u$  from its parent  $G(k)$  we need the probability that individual  $k$  will have exactly the same nucleotide (i.e.,  $+1$  or  $-1$ ) as  $G(k)$ . We assume that the probability of undergoing  $n$  mutations in site  $u$  is Poisson distributed:

$$P^k(n) = \frac{e^{-\mu} \mu^n}{n!}. \quad (8)$$

Each mutation switches the nucleotide (i.e.,  $S_u^k \rightarrow -S_u^k$ ). Then the probability of observing an even number of mutations, so that the nucleotide at site  $u$  does not change the nucleotide is

$$P(S_u^G = S_u^k) = \sum_{i=0}^{\infty} \frac{e^{-\mu} \mu^{2i}}{(2i)!} = e^{-\mu} \sum_{i=0}^{\infty} \frac{\mu^{2i}}{(2i)!} = e^{-\mu} \cosh \mu. \quad (9)$$

The probability of an odd number of mutations, changing the nucleotide, is

$$P(S_u^G = -S_u^k) = \sum_{i=0}^{\infty} \frac{e^{-\mu} \mu^{2i+1}}{(2i+1)!} = e^{-\mu} \sum_{i=0}^{\infty} \frac{\mu^{2i+1}}{(2i+1)!} = e^{-\mu} \sinh \mu. \quad (10)$$

Note that we can have  $0, 1, 2, \dots, n$  mutations in site  $u$  of the new offspring  $k$ , but because the mutation rate,  $\mu$ , is small, most of the probability density is concentrated in the 0 and 1 point mutation cases. The probabilities can be found by solving the system:

$$\begin{cases} (P(S_u^G = S_u^k)) - (P(S_u^G = -S_u^k)) = e^{-\mu} \cosh \mu - e^{-\mu} \sinh \mu = e^{-2\mu} \\ (P(S_u^G = S_u^k)) + (P(S_u^G = -S_u^k)) = \sum_{i=0}^{\infty} \frac{e^{-\mu} \mu^i}{(i)!} = e^{\mu} e^{-\mu} = 1, \end{cases} \quad (11)$$

thus,

$$\begin{cases} P(S_u^G = S_u^k) = \frac{1}{2}(1 + e^{-2\mu}), \\ P(S_u^G = -S_u^k) = \frac{1}{2}(1 - e^{-2\mu}). \end{cases} \quad (12)$$

This derivation is similar to those of Peliti, Serva, Higgs and Derrida [28,76,77], but we consider here nucleotides instead of alleles.

In the asexual model, each individual  $k$  is generated by one parent,  $G(k)$ . The expected fraction of nucleotides in  $k$  shared with each individual  $j$  in the population ( $E[f^{kj}]$ ) is, using equation 12,

$$E[f^{kj}] = [f^{G(k)j}(P(S_u^{G(k)} = S_u^k) + (1 - f^{G(k)j})(P(S_u^{G(k)} = -S_u^k))]. \quad (13)$$

Substituting 12 in 13 we have

$$E[f^{kj}] = \left[ f^{G(k)j} \frac{1}{2}(1 + e^{-2\mu}) + (1 - f^{G(k)j}) \frac{1}{2}(1 - e^{-2\mu}) \right]. \quad (14)$$

Substituting  $f^{G(k)j} = \frac{1 + q^{G(k)j}}{2}$  from equation 7 then gives

$$E[f^{kj}] = \frac{1}{2} \left[ \frac{1 + q^{G(k)j}}{2} + e^{-2\mu} \left( \frac{1 + q^{G(k)j}}{2} \right) + \frac{1 - q^{G(k)j}}{2} - e^{-2\mu} \left( \frac{1 - q^{G(k)j}}{2} \right) \right], \quad (15)$$

and after simplification we obtain

$$E[f^{kj}] = \frac{1}{2} [1 + e^{-2\mu} q^{G(k)j}]. \quad (16)$$

Substituting equation 16 into 6 leads to

$$E[q^{kj}] = 2E[f^{kj}] - 1 = 1 + e^{-2\mu} q^{G(k)j} - 1, \quad (17)$$

and from 17 we get

$$\begin{cases} E[q^{kj}] = e^{-2\mu} q^{G(k)j}, \\ E[q^{kk}] = 1, \end{cases} \quad (18)$$

and equation 3 is derived from this expectation. We used this equation to simulate the mean genetic similarity in the transients,  $\hat{Q}$  and we also used it to calculate the mean genetic similarity of the matrix,  $Q$ , at steady-state for asexual populations,  $Q^* = \frac{1}{(\theta+1)}$ , where  $\theta = 4N_e\mu$  for small  $\mu$  and  $N_e$  is the effective number of individuals in the population [28].

## Simulations

Our simulation is a stochastic, individual-based, zero-sum birth-death model of a sexual and asexual population with overlapping generations. For the simulations reported in the paper, we considered  $N_e$  haploid individuals where only one individual can exist in each site. Simulations were carried out with an initial population,  $N_e = 10^3$  individuals, and this initial population size remained constant throughout the simulations. Results for Figs. 3 and 4 were obtained after  $10^4$  replicates and  $2 \cdot 10^3$  generations of a single model run, where a generation is an update of  $N_e$  time steps. Geographic distance between each pair of individuals  $i$  and  $j$ ,  $d_{ij}$ , was sampled from a normal distribution,  $\mathcal{N}(\hat{d}, \sigma^2)$  and negative values were discarded. Results were qualitatively the same after varying  $\sigma^2 \in \{0.4, 0.1\}$ . In the transients each replicate stops after the mean of the genetic similarity matrix,  $\hat{Q}$ , reached the values,  $\hat{Q} \in \{0.97, 0.95\}$  with all the replicates satisfying  $q^{\min} > \hat{Q}$  (Figs. 3b, 3d and 4a). In the last stage parameter values were chosen to satisfy the mathematical condition required for speciation,  $q^{\min} > Q^*$  [28] (Figs. 3c, 3e and 4b–c). Steady-state was verified by checking the constancy of the mean genetic similarity value during the last  $5 \times 10^2$  generations within each replicate regardless the initial value of  $d^{\max}$ . We explored a broad range of parameter combinations with mutation rate,  $\mu \in \{10^{-3}, 10^{-5}\}$ , the maximum geographic distance for mating and dispersal,  $d^{\max} \in \{1, 0\}$ , and two cut-off values to count species richness in the transient and equilibrium dynamics: the minimum genetic similarity value to define a species in the transients,  $q^{\min} \in \{0.99, 0.97\}$ , and the minimum genetic similarity value to define a species at equilibrium  $q^{\min} \in \{0.97, 0.75\}$ , respectively.

## Acknowledgments

We thank Andrew Allen and Ole Seehausen for useful comments on the development of the present study.

## Author Contributions

Conceived and designed the experiments: CJM SA. Performed the experiments: CJM DA SA. Analyzed the data: CJM DA SA RSC. Contributed reagents/materials/analysis tools: CJM DA SA RSC RSE. Wrote the paper: CJM DA SA RSC RSE.



## References

- Williams GC (1975) Sex and Evolution. Princeton: Princeton University Press. 210 p.
- Maynard Smith J (1978) The Evolution of Sex. Cambridge: Cambridge University Press. 236 p.
- Bell G (1982) The Masterpiece of Nature: The Evolution and Genetics of Sexuality. California: University of California Press. 635 p.
- Misevic D, Kouyos RD, Bonhoeffer S (2009) Predicting the evolution of sex on complex fitness landscapes. *PLoS Comput Biol* 5: e1000510.
- Mayr E (1970) Populations, Species and Evolution. Cambridge: Harvard University Press. 472 p.
- Orr HA (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139: 1805–1813.
- Coyne JA, Orr A (1998) The evolutionary genetics of speciation. *Philos T Roy Soc B* 353: 287–305.
- Barracough TG, Birky CJW, Burt A (2003) Diversification in sexual and asexual organisms. *Evolution* 57: 2166–2172.
- Hillis DM (2007) Asexual evolution: Can species exist without sex? *Curr Biol* 17: R543–R544.
- Fraser C, Hanage WP, Spratt BG (2007) Recombination and the nature of bacterial speciation. *Science* 315: 476–480.
- Fontaneto D, Herniou EA, Boschetti C, Caprioli M, Melone G, et al. (2007) Independently evolving species in asexual bdelloid rotifers. *PLoS Biol* 5: e87.
- Otto S (2009) The evolutionary enigma of sex. *Am Nat* 174: S1–S14.
- Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* 394: 69–72.
- Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* 323: 737–741.
- Butlin R, Bridle J, Schluter D (2009) Speciation and Patterns of Diversity. Cambridge: Cambridge University Press. 346 p.
- Gavrilets S, Losos JB (2009) Adaptive radiation: Contrasting theory with data. *Science* 323: 732–737.
- The Marie Curie SPECIATION Network, Butlin R, Debelles A, Kerth C, Snook RR, et al. (2011) What do we need to know about speciation? *Trends Ecol Evol* 27: 27–39.
- Kimura M, Ohta T (1971) Theoretical Aspects of Population Genetics. Princeton: Princeton University Press. 232 p.
- Kirkpatrick M, Jenkins CD (1989) Genetic segregation and the maintenance of sexual reproduction. *Nature* 339: 300–301.
- Rice WR, Chippindale AK (2001) Sexual recombination and the power of natural selection. *Science* 294: 555–559.
- Butlin R (2002) The cost and benefits of sex: new insights from old asexual lineages. *Nat Rev Genet* 3: 311–317.
- Poon A, Chao L (2004) Drift increases the advantage of sex in RNA bacteriophage Phi 6. *Genetics* 166: 19–24.
- Goddard MR, Godfray HCJ, Burt A (2005) Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* 434: 636–640.
- Livnat A, Papadimitriou C, Dushoff J, Feldman MW (2008) A mixability theory for the role of sex in evolution. *P Natl Acad Sci USA* 105: 19803–19808.
- Kondrashov FA, Kondrashov AS (2001) Multidimensional epistasis and the disadvantage of sex. *P Natl Acad Sci USA* 98: 12089–12092.
- Futuyma DJ (2010) Evolutionary constraint and ecological consequences. *Evolution* 64: 1865–1884.
- Gorelick R, Heng HHQ (2011) Sex reduces genetic variation: A multidisciplinary review. *Evolution* 65: 1088–1098.
- Higgs PG, Derrida B (1992) Genetic distance and species formation in evolving populations. *J Mol Evol* 35: 454–465.
- Manzo F, Peliti L (1994) Geographic speciation in the Higgs-Derrida model of species formation. *J Phys A-Math Gen* 27: 7079–7086.
- Hoelzer GA, Drewes R, Meier J, Doursat R (2008) Isolation-by-distance and outbreeding depression are sufficient to drive parapatric speciation in the absence of environmental fluctuations. *PLoS Comput Biol* 4: e1000126.
- de Aguiar MAM, Baranger M, Baptestini EM, Kaufman L, Bar-Yam Y (2009) Global patterns of speciation and diversity. *Nature* 460: 384–387.
- Melian CJ, Alonso D, Vázquez DP, Regetz J, Allesina S (2010) Frequency-dependent selection predicts patterns of radiations and biodiversity. *PLoS Comput Biol* 6: e1000892.
- Hubbell SP (2001) The Unified Neutral Theory of Biodiversity and Biogeography. Princeton: Princeton University Press. 448 p.
- Chave J, Muller-Landau HC, Levin SA (2002) Comparing classical community models: theoretical consequences for patterns of diversity. *Am Nat* 159: 1–23.
- Allen AP, Savage VM (2007) Setting the absolute tempo of biodiversity dynamics. *Ecol Lett* 10: 637–646.
- Etienne RS, Apol MEF, Oliff H, Weissing FJ (2007) Modes of speciation and the neutral theory of biodiversity. *Oikos* 116: 241–258.
- Mouillot D, Gaston KJ (2007) Geographic range size heritability: what do neutral models with different modes of speciation predict? *Global Ecol Biogeogr* 16: 367–380.
- Leigh EG (2007) Neutral theory: a historical perspective. *J Evolution Biol* 20: 2075–2091.
- Haegeman B, Etienne RS (2009) Neutral models with generalized speciation. *B Math Biol* 71: 1507–1519.
- Beeravolu CR, Couteron P, Pélissier R, Munoz F (2009) Studying ecological communities from a neutral standpoint: A review of models' structure and parameter estimation. *Ecol Model* 220: 2603–2610.
- Rosindell J, Cornell SJ, Hubbell SP, Etienne RS (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol Lett* 13: 716–727.
- O'Dwyer JP, Green JL (2010) Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. *Ecol Lett* 13: 87–95.
- Etienne RS, Haegeman B (2011) The neutral theory of biodiversity with random fission speciation. *Theor Ecol* 4: 87–109.
- Etienne RS, Rosindell J (2011) Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Syst Biol* in press.
- Wiens JJ (2004) What is speciation and how should we study it? *Am Nat* 163: 914–923.
- Rousset F (2004) Genetic Structure and Selection in Subdivided Populations. Princeton: Princeton University Press. 288 p.
- Mallet J (2008) A species definition for the Modern Synthesis. *Trends Ecol Evol* 10: 294–299.
- Wagner A (2008) Neutralism and selectionism: a network-based reconciliation. *Nat Rev Genet* 9: 965–974.
- Bateson W (1909) Heredity and variation in modern lights. In: Seward AC, ed. Darwin and Modern Science. Cambridge: Cambridge University Press. pp 85–101.
- Dobzhansky T (1936) Studies on hybrid sterility. ii. localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* 21: 113–135.
- Muller HJ (1942) Isolating mechanisms, evolution, and temperature. *Biol Symp* 6: 71–125.
- Coyne JA (1992) Genetics and speciation. *Nature* 355: 511–515.
- Coyne JA, Orr HA (2004) Speciation. Massachusetts: Sinauer, Sunderland. 545 p.
- Welch JJ (2004) Accumulating Dobzhansky-Muller incompatibilities: Reconciling theory and data. *Evolution* 58: 1145–1156.
- Kirkpatrick M, Ravigne V (2002) Speciation by natural and sexual selection: models and experiments. *Am Nat* 159: S22–S35.
- Moritz C, Schneider CJ (1992) Evolutionary relationships within the ensatina eschscholtzii complex confirm the ring species interpretation. *Syst Biol* 41: 273–291.
- Gavrilets S (2004) Fitness Landscapes and the Origin of Species. Princeton: Princeton University Press. 432 p.
- Goldschmidt RB (1949) The intersexual males of the beaded minute combination in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 35: 314–316.
- Hubbell SP, Lake J (2003) The Neutral Theory of Biodiversity and Biogeography, and Beyond. In: Blackburn TM, Gaston KJ, eds. Macroecology: patterns and processes. Oxford: Blackwell. pp 45–63.
- Colegrave N (2002) Sex releases the speed limit on evolution. *Nature* 420: 664–666.
- Arjan J, de Visser GM, Elena SF (2007) The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* 8: 139–149.
- Heubel KU, Rankin DJ, Kokko H (2009) How to go extinct by mating too much: Population consequences of male mate choice and efficiency in a sexual-asexual species complex. *Oikos* 118: 513–520.
- Bell G (2010) Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philos T Roy Soc B* 365: 87–97.
- Oren A (2004) Prokaryote diversity and taxonomy: Current status and future challenges. *Philos T Roy Soc B* 359: 623–638.
- Fraser C, Alm EJ, Polz MF, Spratt BG, Hanage WP (2009) The bacterial species challenge: making sense of genetic and ecological diversity. *Science* 323: 741–746.
- Malecot G (1970) The Mathematics of Heredity. San Francisco: W. H. Freeman. 92 p.
- Rosindell J, Wong Y, Etienne RS (2008) Coalescence methods for spatial neutral ecology. *Ecol Inform* 3: 259–271.
- Lynch FAGM (2000) The origins of interspecific genomic incompatibility via gene duplication. *Am Nat* 156: 590–605.
- Lynch M (2007) The Origins of Genome Architecture. Massachusetts: Sinauer Associates. 340 p.
- Ma J, Ratan A, Raney BJ, Suh BB, Miller W, et al. (2008) The infinite sites model of genome evolution. *P Natl Acad Sci USA* 105: 14254–14261.
- Bikard D, Patel D, Le Mette C, Giorgi V, Camilleri C, et al. (2009) Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* 323: 623–626.
- Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, et al. (2010) Species selection maintains self-incompatibility. *Science* 330: 493–495.
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN, ed. Mammalian Protein Metabolism. New York: Academic Press. pp 21–132.
- Kimura M (1983) The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press. 384 p.
- Durrett R (2008) Probability Models for DNA Sequence Evolution. New York: Springer. 432 p.
- Derrida B, Peliti L (1991) Evolution in a flat fitness landscape. *B Math Biol* 53: 355–382.
- Serva M, Peliti L (1991) A statistical model of an evolving population with sexual reproduction. *J Phys A-Math Gen* 24: L705–L709.